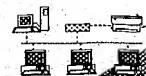BIOTECHNOLOGY
SYSTEMS
BRANCH

# RAW SEQUENCE LISTING
# ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/765,111_

Source: _OIPE_

Date Processed by STIC: _2/2/2001_

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**
**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**
**PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)**
**PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)**

TO **REDUCE** ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

## Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**
**http://www.uspto.gov/web/offices/pac/checker**

OIPE

```
RAW SEQUENCE LISTING              DATE: 02/02/2001
PATENT APPLICATION:  US/09/765,111   TIME: 11:51:20

    Input Set : A:\Seqlist.txt
    Output Set: N:\CRF3\02022001\I765111.raw
```

**Does Not Comply**
**Corrected Diskette Needed**

```
   4 <110> APPLICANT: Fletcher, Jonathan A.
   5        Kroll, Todd G.
   7 <120> TITLE OF INVENTION: PAX8-PPARgamma NUCLEIC ACID MOLECULES
   8        AND POLYPEPTIDES AND USES THEREOF
  11 <130> FILE REFERENCE: B0801/7196 (ERP/MAT)
  13 <140> CURRENT APPLICATION NUMBER: US/09/765,111
  13 <141> CURRENT FILING DATE: 2001-01-18
  13 <150> PRIOR APPLICATION NUMBER: US 60/177,109
  14 <151> PRIOR FILING DATE: 2000-01-20
  16 <150> PRIOR APPLICATION NUMBER: US 60/225,079
  17 <151> PRIOR FILING DATE: 2000-08-14
  19 <160> NUMBER OF SEQ ID NOS: 47
  21 <170> SOFTWARE: FastSEQ for Windows Version 3.0
```

OK

```
ERRORED SEQUENCES

1091 <210> SEQ ID NO: 13
1092 <211> LENGTH: 2711   1372 shown (p.3)
1093 <212> TYPE: DNA
1094 <213> ORGANISM: Homo Sapiens
1096 <220> FEATURE:
1097 <221> NAME/KEY: CDS
1098 <222> LOCATION: (11)...(1363)
1100 <400> SEQUENCE: 13
1101  gaattcggcg atg cct cac aac tcc atc aga tct ggc cat gga ggg ctg        49
1102              Met Pro His Asn Ser Ile Arg Ser Gly His Gly Gly Leu
1103               1               5                  10
1105  aac cag ctg gga ggg gcc ttt gtg aat ggc aga cct ctg ccg gaa gtg       97
1106  Asn Gln Leu Gly Gly Ala Phe Val Asn Gly Arg Pro Leu Pro Glu Val
1107   15                  20                  25
1109  gtc cgc cag cgc atc gta gac ctg gcc cac cag ggt gta agg ccc tgc      145
1110  Val Arg Gln Arg Ile Val Asp Leu Ala His Gln Gly Val Arg Pro Cys
1111   30                  35                  40                  45
1113  gac atc tct cgc cag ctc cgc gtc agc cat ggt tgc gtc agc aag atc      193
1114  Asp Ile Ser Arg Gln Leu Arg Val Ser His Gly Cys Val Ser Lys Ile
1115               50                  55                  60
1117  ctt ggc agg tac tac gag act ggc agc atc cgg cct gga gtg ata ggg      241
1118  Leu Gly Arg Tyr Tyr Glu Thr Gly Ser Ile Arg Pro Gly Val Ile Gly
1119               65                  70                  75
1121  ggc tcc aag ccc aag gtg gcc acc ccc aag gtg gtg gag aag att ggg      289
1122  Gly Ser Lys Pro Lys Val Ala Thr Pro Lys Val Val Glu Lys Ile Gly
1123               80                  85                  90
1125  gac tac aaa cgc cag aac cct acc atg ttt gcc tgg gag atc cga gac      337
1126  Asp Tyr Lys Arg Gln Asn Pro Thr Met Phe Ala Trp Glu Ile Arg Asp
1127   95                  100                 105
1129  cgg ctc ctg gct gag ggc gtc tgt gac aat gac act gtg ccc agt gtc      385
```

```
1130   Arg Leu Leu Ala Glu Gly Val Cys Asp Asn Asp Thr Val Pro Ser Val
1131   110               115               120               125
1133   agc tcc att aat aga atc atc cgg acc aaa gtg cag caa cca ttc aac    433
1134   Ser Ser Ile Asn Arg Ile Ile Arg Thr Lys Val Gln Gln Pro Phe Asn
1135               130               135               140
1137   ctc cct atg gac agc tgc gtg gcc acc aag tcc ctg agt ccc gga cac    481
1138   Leu Pro Met Asp Ser Cys Val Ala Thr Lys Ser Leu Ser Pro Gly His
1139               145               150               155
1141   acg ctg atc ccc agc tca gct gta act ccc ccg gag tca ccc cag tcg    529
1142   Thr Leu Ile Pro Ser Ser Ala Val Thr Pro Pro Glu Ser Pro Gln Ser
1143           160               165               170
1145   gat tcc ctg ggc tcc acc tac tcc atc aat ggg ctc ctg ggc atc gct    577
1146   Asp Ser Leu Gly Ser Thr Tyr Ser Ile Asn Gly Leu Leu Gly Ile Ala
1147           175               180               185
1149   cag cct ggc agc gac aag agg aaa atg gat gac agt gat cag gat agc    625
1150   Gln Pro Gly Ser Asp Lys Arg Lys Met Asp Asp Ser Asp Gln Asp Ser
1151   190               195               200               205
1153   tgc cga cta agc att gac tca cag agc agc agc agc gga ccc cga aag    673
1154   Cys Arg Leu Ser Ile Asp Ser Gln Ser Ser Ser Ser Gly Pro Arg Lys
1155               210               215               220
1157   cac ctt cgc acg gat gcc ttc agc cag cac cac ctc gag ccg ctc gag    721
1158   His Leu Arg Thr Asp Ala Phe Ser Gln His His Leu Glu Pro Leu Glu
1159               225               230               235
1161   tgc cca ttt gag cgg cag cac tac cca gag gcc tat gcc tcc ccc agc    769
1162   Cys Pro Phe Glu Arg Gln His Tyr Pro Glu Ala Tyr Ala Ser Pro Ser
1163               240               245               250
1165   cac acc aaa ggc gag cag ggc ctc tac ccg ctg ccc ttg ctc aac agc    817
1166   His Thr Lys Gly Glu Gln Gly Leu Tyr Pro Leu Pro Leu Leu Asn Ser
1167           255               260               265
1169   acc ctg gac gac ggg aag gcc acc ctg acc cct tcc aac acg cca ctg    865
1170   Thr Leu Asp Asp Gly Lys Ala Thr Leu Thr Pro Ser Asn Thr Pro Leu
1171   270               275               280               285
1173   ggg cgc aac ctc tcg act cac cag acc tac ccc gtg gtg gca gat cct    913
1174   Gly Arg Asn Leu Ser Thr His Gln Thr Tyr Pro Val Val Ala Asp Pro
1175               290               295               300
1177   cac tca ccc ttc gcc ata aag cag gaa acc ccc gag gtg tcc agt tct    961
1178   His Ser Pro Phe Ala Ile Lys Gln Glu Thr Pro Glu Val Ser Ser Ser
1179               305               310               315
1181   agc tcc acc cct tcc tct tta tct agc tcc gcc ttt ttg gat ctg cag    1009
1182   Ser Ser Thr Pro Ser Ser Leu Ser Ser Ser Ala Phe Leu Asp Leu Gln
1183               320               325               330
1185   caa gtc ggc tcc ggg gtc ccg ccc ttc aat gcc ttt ccc cat gct gcc    1057
1186   Gln Val Gly Ser Gly Val Pro Pro Phe Asn Ala Phe Pro His Ala Ala
1187           335               340               345
1189   tcc gtg tac ggg cag ttc acg ggc cag gcc ctc ctc tca ggg cga gag    1105
1190   Ser Val Tyr Gly Gln Phe Thr Gly Gln Ala Leu Leu Ser Gly Arg Glu
1191   350               355               360               365
1193   atg gtg ggg ccc acg ctg ccc gga tac cca ccc cac atc ccc acc agc    1153
1194   Met Val Gly Pro Thr Leu Pro Gly Tyr Pro Pro His Ile Pro Thr Ser
```

RAW SEQUENCE LISTING                          DATE: 02/02/2001
PATENT APPLICATION:  US/09/765,111            TIME: 11:51:20

Input Set : A:\Seqlist.txt
Output Set: N:\CRF3\02022001\I765111.raw

```
1195                    370              375              380
1197  gga cag ggc agc tat gcc tcc tct gcc atc gca ggc atg gtg gca gga    1201
1198  Gly Gln Gly Ser Tyr Ala Ser Ser Ala Ile Ala Gly Met Val Ala Gly
1199                385              390              395
1201  agt gaa tac tct ggc aat gcc tat ggc cac acc ccc tac tcc tcc tac    1249
1202  Ser Glu Tyr Ser Gly Asn Ala Tyr Gly His Thr Pro Tyr Ser Ser Tyr
1203            400              405              410
1205  agc gag gcc tgg cgc ttc ccc aac tcc agc ttg ctg agt tcc cca tat    1297
1206  Ser Glu Ala Trp Arg Phe Pro Asn Ser Ser Leu Leu Ser Ser Pro Tyr
1207        415              420              425
1209  tat tac agt tcc aca tca agg ccg agt gca ccg ccc acc act gcc acg    1345
1210  Tyr Tyr Ser Ser Thr Ser Arg Pro Ser Ala Pro Pro Thr Thr Ala Thr
1211  430              435              440              445
1213  gcc ttt gac cat ctg tag ttgaagctt                                 1372
1214  Ala Phe Asp His Leu  *
E--> 1215                 450
```

FYI!

**Please Note:**
**Use of n and/or Xaa have been detected in the Sequence Listing.  Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.**

VERIFICATION SUMMARY                         DATE: 02/02/2001
PATENT APPLICATION:  US/09/765,111           TIME: 11:51:21

Input Set : A:\Seqlist.txt
Output Set: N:\CRF3\02022001\I765111.raw

```
L:13   M:270 C: Current Application Number differs, Replaced Current Application No
L:13   M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1215 M:252 E: No. of Seq. differs, <211>LENGTH:Input:2711 Found:1372 SEQ:13
L:2364 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:36
L:2364 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:36
L:2364 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:36
L:2386 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:37
L:2386 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:37
L:2386 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:37
L:2402 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:38
L:2402 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:38
L:2402 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:38
L:2418 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:39
L:2418 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:39
L:2418 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:39
L:2452 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:40
L:2452 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:40
L:2452 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:40
L:2478 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:41
L:2478 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:41
L:2478 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:41
L:2501 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:2502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:2515 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:43
L:2515 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:43
L:2515 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:43
L:2530 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44
L:2531 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44
L:2544 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:45
L:2544 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:45
L:2544 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:45
L:2559 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:2560 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:2573 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:47
L:2573 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:47
L:2573 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:47
```